Neighborhood dynamics and price effects of Superfund site clean-up

Douglas S. Noonan
Assistant Professor
Georgia Institute of Technology
douglas.noonan@pubpolicy.gatech.edu

Douglas J. Krupka
Senior Research Associate
IZA, Bonn
krupka@iza.org

Brett M. Baden
Senior Associate
Industrial Economics, Inc.
bbaden@indecon.com

JEL classification: Q51, Q53, Q58, R21

[Preprint; final version available as:

Noonan, Douglas S., Douglas J. Krupka, and Brett M. Baden. 2007. "Neighborhood Dynamics and Price Effects of Superfund Site Clean-up." *Journal of Regional Science* 47(4): 665-692.]

ABSTRACT

Numerous hedonic price analyses estimate price effects associated with hazardous waste site remediation or other environmental variation. This paper estimates a neighborhood transition model to capture the direct price effect from Superfund site cleanup and the indirect price effects arising from residential sorting and changes in investment in the housing stock following cleanup. First-difference models of neighborhood change and a national sample are used. This approach fails to find consistent positive direct price effects. Positive indirect effects, however, may arise through residential sorting and neighborhood investment spurred by remediation. The findings can be sensitive to policy endogeneity and model specification.

* We would like to thank Matthew Kahn, Geoffrey Turnbull, Jeffrey Zabel, and two anonymous referees for their helpful comments. "This document contains demographic data from GeoLytics, E. Brunswick, NJ." This material is based upon work supported by the National Science Foundation under Grant No. 0433165.

I. INTRODUCTION

Hedonic price analysis (Rosen 1974) is frequently used to estimate the implicit price of structural or neighborhood characteristics. It has often been applied to environmental goods such as proximity to hazardous sites. There is a temptation to use the coefficients identified in the first-stage price regression from cross-sectional variation to predict within-unit price changes associated with policy shocks. Yet, even if important concerns about unobservables are addressed, these hedonic estimates offer limited insight into how the change in environmental quality affects markets. Other (nonprice) impacts on neighborhood composition are often overlooked.

This paper proposes to extend the standard hedonic approach by tracing the pathways through which environmental change can affect price. For prices to change, the new equilibrium requires some turnover in the housing market and, most likely, other shifts in neighborhood composition and characteristics of the housing stock. Neighborhood sorting occurs and reinvestment takes place as hazardous waste sites are cleaned up. High demanders may bid up prices following improved environmental quality, yet the ensuing change in residents (and their houses, stores, etc.) will in turn affect observed prices. If neighborhood characteristics affect price, the effect of clean-ups on price *through* neighborhood transition and reinvestment will be important parts of the total effect of hazardous waste clean-ups. Conventional studies that neglect the effect of environmental quality on neighborhood composition and investment decisions may miss these effects.

This paper uses panel data to estimate a system of equations that allows for endogeneity among prices, neighborhood characteristics, housing stock variables, and environmental quality. Although the data are not ideal, our estimates of the direct effect of hazardous waste clean-ups resemble those estimated in other studies using sales data. However, our system of equations estimates allow us to compute the indirect effects as well, which improves our understanding of expected price changes. We find some of these indirect effects to be substantively significant.

The implications of these results are threefold. First, the presence of substantial indirect effects of changes in environmental quality should guide our interpretation of existing and future empirical estimates of price effects. Second, the relative magnitude of the indirect effects suggests that these effects should be considered carefully in cost-benefit analyses of hazardous waste clean-ups, or other policy interventions that might cause neighborhood transition. Third, that sorting does appear to occur because of clean-up activity has implications for the environmental justice literature. Cross-sectional models of minority group exposure may not be adequate given these dynamic responses of housing markets to environmental improvements.

The rest of the paper runs as follows: Section II briefly reviews the relevant literatures. Section III lays out the empirical model, derives the total effect of a clean-up, and describes the data. Section IV presents the results, and Section V concludes.

II. ANALYZNG THE EFFECTS OF SUPERFUND SITES

The hedonic literature concerning price effects of Superfund NPL sites is sizable. Kiel and Williams (2007) provide a recent review. While some studies, such as Greenberg and Hughes (1992) use simple means comparisons to draw inferences about the effects of environmental hazards on property values, most economic studies estimate price effects from first-stage hedonic regressions. A review of at almost a dozen studies[1] that estimate comparable price effects for proximity to hazardous waste sites reveals some general tendencies. Most include neighborhood-level demographic controls, while a few (e.g., Chattopadhyay et al. 2005, Kiel 1995) do not. Six of these studies report adjacency effects, which range from -12% to 1% of the total property value. The rest of the papers estimate price gradients around these sites. These gradients range from insignificant to about six percent per mile away from the hazardous waste site, with most effects deemed insignificant between one and six miles from the site. Ketkar (1992) used aggregated census data for the dependent variable. In that New Jersey sample, hazardous waste sites account for 2% lower median housing values.

The typical approach in this literature is to use results of a first-stage hedonic regression, where the key variable is some "distance to site," often interacted with information about or clean-up status of the site. This conventional research design essentially identifies changes in property values by comparing prices near the site with prices elsewhere (or by comparing prices before and after a change in status). Greenstone and Gallagher (2005) criticize this model specification as being particularly susceptible to omitted variable bias, especially using cross-sectional variation in prices and proximity to Superfund sites. The present paper seeks to address Greenstone and Gallagher's concerns through a panel structure and other efforts to mitigate omitted variable bias while also exploring nonprice effects of cleanups. Recent work by Cameron and McConnaha (2006), Banzhaf and Walsh (2006), and Bayer et al. (2006) have all pointed to the important role of environmentally-induced migration. If sorting occurs in housing markets, conventional hedonic models that include neighborhood composition variables risk including controls that are jointly endogenous. Consistent estimates of price (and nonprice) impacts of environmental shocks involves careful attention to both unobservables and the interdependence of price and neighborhood dynamics.

The co-location of NPL sites and residents is a major policy issue, especially in terms of the equity of exposure. The environmental justice movement and literature have often focused on hazardous waste facilities. To varying degrees, assessments of distributional equity have accounted for migration and sorting in response to siting decisions. Been (1997), Anderton et al. (1994), and Baden and Coursey (2002) all cast doubt on the hypothesis that siting follows race. Alternatively, Hamilton (1995) finds facility expansions to follow neighborhoods where collective political action was weakest. Gayer (2000) finds likewise for levels of risks, while Shapiro (2005) finds similarly for changes in risk levels. One implication is that the EPA may base decisions about cleanup, in part, on the political influence of a community. Another implication is that surprisingly little work exists in the environmental justice literature that explicitly measures a dynamic model where households migrate and environmental quality improves. Most of these studies examine sorting around existing sites. Far less effort has gone into rigorously investigating neighborhood transition in the wake of Superfund site remediation. If there are price effects associated with environmental remediation, then neighborhood sorting should follow. While this dynamic has yet to be observed following siting, at least in the environmental justice literature (Ringquist 2006), we offer new evidence on sorting following remediation. Moreover, shifts in rentership rates are also estimated, with implications for whether residents can capture subsequent property value gains (see discussion in Sieg et al. 2004).

III. THEORY AND EMPIRICAL METHODOLOGY

A. Theoretical model

In general, hedonic studies use cross-sectional data to estimate a first-stage equation of the form:

(1) $$P_{it} = \beta_{0t} + \beta_E E_{it} + \beta_S S_{it} + \beta_N N_{it} + \beta_M M_{it} + \beta_G G_i + \varepsilon_{1it}$$

where t indexes time, i indexes individual housing units, P is the house value, E measures environmental quality (which is a negative function of the presence of a polluted site), S is a vector of structural characteristics of a property, N is a set of neighborhood demographic characteristics, M is a vector of characteristics of the municipality that may vary over time and also affect prices, and G is a set of time-invariant characteristics that affect price (such as distance to the CBD). The hedonic price, β_E, is typically taken to also represent the willingness to pay – by the marginal consumer in that housing market – for a marginal increase in environmental quality.

One potential problem with a simple OLS approach to the hedonic equation in levels is that some of the components of G will be unobserved and correlated with the other variables of interest. To mitigate this problem, we estimate the model in first differences (Mendelsohn et al. 1992, Zabel 1999). This strategy purges our parameter estimate of bias from the omission of time-invariant variables, and we thus identify the parameters from within-observation changes in environmental quality, neighborhood demographic conditions, and structural characteristics, as in Equation (2):

(2) $$\dot{P}_{it} = \beta_0 + \beta_E \dot{E}_{it} + \beta_S \dot{S}_{it} + \beta_N \dot{N}_{it} + \beta_M \dot{M}_{it} + \dot{\varepsilon}_{1it}$$

where $\dot{X}_{it} = X_{i,t} - X_{i,t-1}$ for any variable X.

In this paper, we consider the possibility that Equation (2) is part of a larger system in which many of the key variables are set simultaneously. Estimating the system that simultaneously determines prices and housing and demographic characteristics reveals the direct price effects of environmental clean-ups and enables us to map the pathways through which the indirect effects arise. Our model of structural characteristics explains observed levels of *S* by the lagged level of *S* and the other variables in the system:

(3) $$S_{it} = \gamma_S S_{it-1} + \gamma_{0t} + \gamma_E E_{it} + \gamma_N N_{it} + \gamma_M M_{it} + \gamma_G G_i + \varepsilon_{2it}$$

Here, the housing stock depends on its past levels and environmental quality, neighborhood demographics, and other considerations. Again, taking first-differences to control for time-invariant unobservables yields:

(4) $$\dot{S}_{it} = S_{it} - S_{it-1} = \gamma_S \dot{S}_{it-1} + \gamma_0 + \gamma_E \dot{E}_{it} + \gamma_N \dot{N}_{it} + \gamma_M \dot{M}_{it} + \dot{\varepsilon}_{2it}$$

Surely, all of these adjustments will be gradual in the aggregate, as the existing stock, built before the changes occurred, will not instantaneously be demolished and rebuilt at the new equilibrium specifications.[2] That is, if there are no changes in environmental quality, neighborhood demographics, or other considerations, the housing stock in an area will continue its path towards long-run equilibrium. However, if environmental quality suddenly improves, it might cause people to change the kinds of housing they build.[3] Likewise, neighborhood demographics such as family size and income will also affect the equilibrium quantity and quality of the housing stock if the demand for housing is related to these demographics.

A similar argument holds for neighborhood demographic characteristics. Let the observed neighborhood demographics be explained by:

(5) $$N_{it} = \delta_N N_{it-1} + \delta_{0t} + \delta_E E_{it} + \delta_S S_{it} + \delta_M M_{it} + \delta_G G_i + \varepsilon_{3it}$$

Taking first-differences yields:

(6) $$\dot{N}_{it} = N_{it} - N_{it-1} = \delta_N \dot{N}_{it-1} + \delta_0 + \delta_E \dot{E}_{it} + \delta_S \dot{S}_{it} + \delta_M \dot{M}_{it} + \dot{\varepsilon}_{3it}$$

Thus, *N* follows a partial adjustment process, where changes in environmental quality, prices, structural characteristics, and other factors all explain the observed changes in neighborhood demographics. Demographic groups' differing demands for *E* may cause them to sort into neighborhoods according to their willingness to pay for these attributes (Diamond and Tolley 1982). Similarly, changes in housing stock may attract different types of residents, at least when the capital stock is somewhat inelastic.

Rearranging terms and rewriting the first-differenced system in matrix notation yields Equation (7).

(7) $$\begin{bmatrix} 1 & -\beta_S & -\beta_N \\ 0 & 1 & -\gamma_N \\ 0 & -\delta_S & 1 \end{bmatrix} \begin{bmatrix} \dot{P} \\ \dot{S} \\ \dot{N} \end{bmatrix} = \begin{bmatrix} \beta_E \\ \gamma_E \\ \delta_E \end{bmatrix} \dot{E} + \begin{bmatrix} 0 \\ \gamma_S \dot{S}_{t-1} \\ \delta_N \dot{N}_{t-1} \end{bmatrix} + \begin{bmatrix} \beta_M \\ \gamma_M \\ \delta_M \end{bmatrix} \dot{M} + \begin{bmatrix} \dot{\varepsilon}_1 \\ \dot{\varepsilon}_2 \\ \dot{\varepsilon}_3 \end{bmatrix}$$

where $\dot{X} = X_t - X_{t-1}$ for any variable *X*.

In this paper, we are specifically interested in the effects of $\dot{E}$, especially when *E* changes due to policy intervention, as in the case of Superfund site clean-ups. Given the system of Equations (7), the total effect of a clean-up ($\dot{E}$) can be seen to depend not solely on its direct effect (β_E), but also on its indirect effects. Totally differentiating and dividing through by $d\dot{E}$, while recognizing the lagged differences in *P*, *S*, and *N* will not depend on $\dot{E}$, yields:

$$\begin{bmatrix} 1 & -\beta_S & -\beta_N \\ 0 & 1 & -\gamma_N \\ 0 & -\delta_S & 1 \end{bmatrix} \begin{bmatrix} d\dot{P}/d\dot{E} \\ d\dot{S}/d\dot{E} \\ d\dot{N}/d\dot{E} \end{bmatrix} = \begin{bmatrix} \beta_E \\ \gamma_E \\ \delta_E \end{bmatrix}$$

We can then use Cramer's Rule to obtain the total effect of a change in *E*:

$$\text{(8)} \qquad \frac{d\dot{P}}{d\dot{E}} = \frac{\beta_E + \beta_S\gamma_E + \beta_N\delta_E + \beta_S\gamma_N\delta_E + \beta_N\delta_S\gamma_E - \beta_E\gamma_N\delta_S}{1-\gamma_N\delta_S}$$

The first term in the numerator is the direct effect on price. The second and third terms in the numerator are the first-order indirect effect: $\dot{E}$'s effect on $\dot{P}$ through $\dot{S}$ and $\dot{N}$. The third and fourth terms are the second-order indirect effects: $\dot{E}$'s effect on $\dot{P}$ through $\dot{S}$'s effect on $\dot{N}$ and $\dot{N}$'s effect on $\dot{S}$. The final term corrects for double counting. The denominator is a sort of multiplier effect. If there is no endogeneity in Equation (7),[4] this total derivative reduces to the first three terms in the numerator.

In this application, the system of equations is considerably more complex because *S*, *N*, and *M* are vectors of many variables. Hence, in Equation (3), we assume that each variable in *S* depends on its own lag, the vectors *E*, *N*, *M*, and *G*, *and* the contemporaneous values of the other variables in *S*. Likewise, in Equation (5), each *N* variable depends on its own lag, the vectors *E*, *S*, *M*, and *G*, *and* the contemporaneous values of the other variables in *N*. The system in Equation (7) thus has each $\dot{S}$ and $\dot{N}$ equation dependent on that variable's own lagged difference, $\dot{E}$, the other variables in the $\dot{S}$ and $\dot{N}$ vectors, and additional municipal-level controls in the exogenous $\dot{M}$ vector. The time-invariant vector of geographic controls, *G*, drops entirely out of the system when first-differenced, assuming time-invariant parameters β_G, γ_G, and δ_G.

An alternative model of the system might allow for property values to enter into Equations (3) and (5) directly. Thus, they appear as:

(3a) $$S_{it} = \gamma_S S_{it-1} + \gamma_{0t} + \gamma_E E_{it} + \gamma_N N_{it} + \gamma_P P_{it} + \gamma_M M_{it} + \gamma_G G_i + \varepsilon_{2it}.$$

(5a) $$N_{it} = \delta_N N_{it-1} + \delta_{0t} + \delta_E E_{it} + \delta_S S_{it} + \delta_P P_{it} + \delta_M M_{it} + \delta_G G_i + \varepsilon_{3it}.$$

In Equation (3a), the substitution towards different types of housing when property prices rise suggests an important role of *P* in explaining *S*. Similar arguments hold for the inclusion of *P* in Equation (5a): higher property values may attract different types of residents. First-differencing these equations completes the alternative system.

(7a) $$\begin{bmatrix} 1 & -\beta_S & -\beta_N \\ -\gamma_P & 1 & -\gamma_N \\ -\delta_P & -\delta_S & 1 \end{bmatrix} \begin{bmatrix} \dot{P} \\ \dot{S} \\ \dot{N} \end{bmatrix} = \begin{bmatrix} \beta_E \\ \gamma_E \\ \delta_E \end{bmatrix} \dot{E} + \begin{bmatrix} 0 \\ \gamma_S \dot{S}_{t-1} \\ \delta_N \dot{N}_{t-1} \end{bmatrix} + \begin{bmatrix} \beta_M \\ \gamma_M \\ \delta_M \end{bmatrix} \dot{M} + \begin{bmatrix} \dot{\varepsilon}_1 \\ \dot{\varepsilon}_2 \\ \dot{\varepsilon}_3 \end{bmatrix}.$$ [5]

In this alternative model, as above, each $\dot{S}$ and $\dot{N}$ equation depends on that variable's own lagged difference, $\dot{E}$, $\dot{P}$, the exogenous $\dot{M}$ vector, and the other variables in the $\dot{N}$ and $\dot{S}$ vectors, respectively.

B. Estimation approach

To identify the parameters in Equation (8), we estimate the system of Equations (7). In this framework, the preferred data set would include a national sample[6] of properties and a rich set of housing and resident characteristics over time. The two most obvious candidates (the American Housing Survey and the Public Use Micro Sample) only provide geographic information at the county-level. Since the effects of hazardous waste sites have been found to be highly localized (Hite et al. 2001, Mendelsohn et al. 1992), such large geographic scales are inadequate for our purposes.

In the absence of national microdata, we use aggregate measures of housing and population characteristics at the neighborhood (block group) level. Using block-group averages and medians, we wish to see how neighborhood transitions induced by site clean-ups affect total changes in prices. There are some advantages to this level of aggregation (Goodman 1977). Coulton et al. (2004) show that the block group matches survey respondents' perceptions of "neighborhood" better than other available levels of aggregation. We use U.S. census data from 1980, 1990, and 2000, processed by Geolytics, Inc. so that block-group boundaries do not change from decade to decade. This geographic consistency across years enables panel data analysis. We treat block groups, the smallest level of aggregation for which our data are available, as the unit of analysis in the first-difference approach.

The use of aggregated data, even at the neighborhood level, limits our ability to infer price effects at the individual level. Nonetheless, some hedonic research has shown that estimates using aggregate data produce reasonably accurate results (Freeman 1979, Nelson 1979, O'Byrne et al. 1985).[7] Moreover, the median housing value in a neighborhood is of considerable policy import. Learning more about the effects of clean-ups on this neighborhood measure is informative, even if it does not recover the true underlying hedonic price. The results based on such aggregate measures can be viewed in an epidemiological light: the effects of average exposure on average outcomes, while not the ideal, are nonetheless interesting.

The estimation strategy employed here attempts to avoid two kinds of bias that would result from estimating Equations (1), (3), and (5) with OLS. The first arises from time-invariant omitted variables. If we had no time-invariant omitted variables and good instruments, then estimating Equations (1), (3), and (5) in levels would be sufficient and straightforward. Yet, to address the serious concerns about unobservable individual-level effects and to help with our search for instruments, we estimate the system in first-differences. This recovers the same set of parameters as found in the system in levels by relying on observations of how these variables respond to changes in the environmental good (something that would not be possible with cross-sectional data alone).

Another source of bias in all three equations stems from endogeneity in the system.[8] Equation (4) has changes in structural characteristics depending on changes in neighborhood demographics just as Equation (6) has demographic trends following from changes in the housing stock. Equation (7a) states that appreciation rates are affected by changes in the structural characteristics while also stating that those changes depend in part on the appreciation rate. To correct for this endogeneity, we estimate Equation (7) as a system of simultaneous equations. Having differenced out all time-invariant determinants of *P*, *S*, and *N*, the search for instruments is made somewhat less arduous. With a few exceptions, we use twice-lagged levels of each variable as instruments.

The consistency of our estimator depends on the validity of our instruments. Under the assumption of white noise error terms in (7), the twice-lagged levels of each variable serve as valid instruments. This follows the recommendations of Arellano (1989). Thus, while $\dot{\varepsilon}_t$ and $\dot{\varepsilon}_{t-1}$ must be uncorrelated, this implies ε_t and ε_{t-2} are uncorrelated for each equation in the system in (7). Sargan tests of overidentification and Durbin-Wu-Hausman tests of endogeneity – both of which are performed on the 2SLS estimation of each equation in the system – serve as diagnostic checks for the validity of these assumptions with these data.

C. Variables and Descriptive statistics

Data from several sources are combined to estimate the model. The results are presented in section IV, emphasizing the estimation of the $\dot{P}$ equation in (4). $\dot{P}$ is the change in the block group's log of median house value from 1990 to 2000.

Our variable of interested is $\dot{E}$, which represents EPA clean-up activity over the 1990's. Derived from public EPA data (EPA 2003), this variable equals one if a block group contains a site that was deleted or partially deleted from the NPL during the 1990s. This is the most complete and final designation of a hazardous waste site, indicating that the EPA is satisfied that the site has been cleaned enough to pose no further health risk. This change in status, more than mere listing or incomplete remediation, should represent improved environmental quality (Greenstone and Gallagher 2005). Because the policy variable, $\dot{E}$, may not be exogenous, we also estimate models with an instrumented $\dot{E}$ (discussed in detail in section D below) to assess the sensitivity of the results.

$\dot{S}$ is a vector of housing characteristics expected to affect prices at an individual as well as an aggregate level. $\dot{S}$ includes changes in eight variables: median year built of housing units, average number of rooms per unit, percent of housing units with gas or electric heating, housing density (housing units per square mile), percent of units in small buildings (containing four or fewer housing units), percent of housing units with complete plumbing, average number of bedrooms, and the percent of housing units that are stand-alone.

Neighborhood demographic characteristics ($\dot{N}$) include the changes in the following eleven variables: log of the neighborhood median household income; percent population that is white but not Hispanic; percent population aged 25 or older who have completed at least a bachelor's degree; percent population below 1.5 times the poverty level; percent of the population employed in manufacturing, warehousing, transportation or utilities industries; percent renter-occupied housing, percent population aged under 18 years; average commute time for people working outside the home; percent of households who do not have a vehicle available; population density; and average people per housing unit.

The *M* vector captures the conditions of the municipality or Census "place". Ideally, $\dot{M}$ would measure changes in important variables like school quality, crime, and public finance attributes of the area. Such variables are unavailable, however, for the nationwide sample and appropriate dates (1980-2000) needed in this analysis. Proxies are constructed using first-differences in the following place-level variables: number of households, median housing value, median rent, median household income, percent of households with children under 18, and percent of households that are married families with kids. This approach groups observations not in an MSA into a single rural group to compute the group-level trends.

Time-invariant components of *G* will cancel out in the first-difference estimation. If we relax the assumption of constant hedonic prices for these characteristics, however, geographic variables may re-enter the model. The β_G would then reflect the change in the hedonic price from 1990 – 2000. The same applies for the other equations for structural and neighborhood characteristics. To account for possible changing influence of unobserved MSA-level characteristics in the various equations in the system, all models are estimated with MSA-level fixed effects. By subtracting the MSA-level means from each variable, the models effectively control for changing prices (and other levels of *S* and *N*) at the metropolitan level over the 1990s. Relaxing the assumption of time-invariant prices also allows a richer set of geographic controls, *G*, to enter the price equation in some specifications.[9] The *G* vector includes a natural amenity index computed at the county level by the USDA ERS (USDA 1999) and a set of interactions between MSA dummies and distance to CBD, which was derived from various Census TIGER files and the National Atlas of the United States (2004). By including these time-invariant factors in our price model, it allows housing price trends to vary according to climate and topography, and across and within MSAs. Table 1 presents the variable names, descriptions and descriptive statistics of all the variables described above.

Table 1: Variable Names, descriptions and descriptive statistics.

Vector	Name	Description[a]	Mean	Standard Deviation
$\dot{P}$	Price	Difference in log of median value, owner-occupied housing	0.3589	0.322
$\dot{E}$	Clean-up, own/adjacent	Own or adjacent block group has an NPL site deleted from list	0.0053	0.073
E	NPL in 1990, own/adjacent	Own or adjacent block group had an NPL site as of 1990	0.0544	0.227
$\dot{S}$	Year built	Difference in median year structure was built	3.3446	20.190
	Rooms	Difference in average number of rooms in housing units	0.1058	0.457
	Utility heat	Difference in percent housing units with gas or electric heat	0.0492	0.102
	Housing density	Difference in housing units per mile^2	52.8743	702.253
	Small structures	Difference in percent housing units sharing structure with 4 or less housing units	0.0004	0.080
	Plumbing	Difference in percent housing units w/ complete plumbing	-0.0007	0.028
	Bedrooms	Difference in average number of bedrooms in housing units	0.0071	0.243
	Solo unit	Difference in percent housing units not sharing structure with any other housing units	0.0031	0.091
$\dot{N}$	Income	Difference in log of median household income	0.3343	0.231
	White	Difference in percent non-Hispanic white population	-0.0678	0.110
	College	Difference in percent population age 25+ with at least college degree	0.0486	0.079

Vector	Name	Description[a]	Mean	Standard Deviation
	Poor	Difference in percent population with income under 1.5 poverty line	-0.0031	0.099
	Blue collar	Difference in percent workers employed in "industrial" sectors	-0.0622	0.087
	Renter	Difference in percent occupied housing units that are renter-occupied	-0.0053	0.096
	Children	Difference in percent population aged 18 or younger	-0.0023	0.062
	Commute	Difference in average travel time for those working outside of home	2.092	4.943
	No vehicle	Difference in percent of households with no vehicle available	-0.0056	0.067
	Population density	Difference in people per mile^2	199.2871	2117.182
	Household size	Difference in people per housing unit	-0.0099	4.966
$\dot{M}$	PlaceHouseholds	Difference in place-level number of households	18019.47	42207.37
	PlaceValue	Difference in place-level median housing value	34706.75	34466.24
	PlaceRent	Difference in place-level median rent	41.9863	80.549
	PlaceIncome	Difference in place-level median household income	10366.28	4852.253
	PlaceKids	Difference in place-level percent of households with children aged 18 or less	-0.0095	0.029
	PlaceFamilies	Difference in place-level percent of households that are married with children	-0.0240	0.028
G	Natural amenities scale	county-level amenity index (composed of topography, temperatures, humidity, and sunlight)	1.0601	3.209
	Distance	log of distance to historic city center	2.8742	0.818
	MSA × Distance	MSA-specific log of distance to historic city center		

[a] All variables are measured as changes from 1990 to 2000, except for E and level variables in G. Even though variables in Table 1 appear in their raw form, all models are estimated after subtracting MSA-level averages from all variables.

To identify the system in Equation (7), numerous instruments are needed. The system has 20 equations (for $\dot{P}$, $\dot{S}$, $\dot{N}$) and 39 endogenous variables when the lagged differences of S and N are also included. In each equation, however, there are only 20 endogenous variables: the 19 endogenous differences and one endogenous lagged difference. The system includes exogenous $\dot{M}$ as regressors in each equation. Excluded instruments include the twice-lagged levels of all P, S, and N variables.[10] The three-stage least squares (3SLS) approach estimates all endogenous variables using the $\dot{E}$, P_{t-2}, S_{t-2},

N_{t-2}, and $\dot{M}_t$ as independent variables. Then, the instrumented versions of the endogenous variables are used to estimate the equations in the system simultaneously, allowing for across-equation correlations in the errors. The 20 exclusion restrictions (twice-lagged levels of *P*, *S*, and *N*) overidentify each of the equations' 19 endogenous regressors (the $\dot{S}$ and $\dot{N}$ vectors, except for the dependent variable being replaced by its lagged difference). A Sargan test of overidentification (with one degree of freedom) offers a specification test for each of the equations.

The alternative model (equation 7a) differs slightly in that each $\dot{S}$ and $\dot{N}$ equation also includes an endogenous $\dot{P}$ regressor. It uses an identical instrument set. The exclusion restrictions (20 twice-lagged levels from $\dot{P}$, $\dot{S}$, and $\dot{N}$) just identify the endogenous variables in each of the $\dot{S}$ and $\dot{N}$ equations (where all of the $\dot{P}$, $\dot{S}$, and $\dot{N}$ variables appear in each, except for the dependent variable being replaced by its lagged difference) and overidentify the price equation (because the lagged difference of price does not appear). A Sargan test is available for the price equation.

D. Endogeneity Issues with $\dot{E}$

Up to this point, we have maintained the assumption that while *P*, *S*, and *N* are determined simultaneously, EPA clean-ups are determined exogenously (i.e., $\dot{E}$ is uncorrelated with $\dot{\varepsilon}_i$ in equations 7 or 7a). This assumption may not hold. Empirically, Viscusi and Hamilton (1999) show that the clean-up standard chosen by the EPA depends on some demographic characteristics of the area surrounding the site. On the other hand, Hird (1993, 1994) shows that the demographics of the county which contains a site has little effect on the progress of a site through remediation. Gupta et al. (1996) found similar results: that clean up decisions for NPL sites were more related to clean up costs and risk factors than to neighborhood demographics. More recently, Daley and Layton (2004) predict site progress through remediation and find no significant effects of area demographics. These studies offer little evidence that deletion, our variable of interest, hinges much on unexplained changes in P, S, and N.

Because all of these studies are done conditional on the presence of a hazardous waste site, the exogeneity of clean-ups among sites might not generalize to our more inclusive sample. Even in our first-differencing context, this should concern us because neighborhoods experiencing clean-ups will be neighborhoods that contained sites in 1990 (i.e., E_{t-1} and $\dot{E}$ are highly correlated). If neighborhoods with Superfund sites in 1990 tended to have systematically higher or lower residual appreciation rates, the β_E could be biased. To control for this possibility, we will also estimate models on the subset of neighborhoods that contained sites in 1990. In this subsample, we were unable to reject the null hypothesis of exogeneity of $\dot{E}$ in most of our equations using a Durbin-Wu-Hausman test. Thus, we feel it is safe to take clean-up as exogenous, conditional on the presence of a site. We report results from models run on the full sample with (Models 2 and 4) and without (Models 1 and 3) instrumenting for $\dot{E}$. We also report comparable results for models run on the subsample of observations with sites as of 1990 (Models 1' and 3').

Along the lines of Gayer (2000), we consider an instrumented version of $\dot{E}$ in Models 2 and 4. $\hat{\dot{E}}$ is estimated from a probit of P_{t-2}, S_{t-2}, N_{t-2}, E_{t-1}, site-specific characteristics, and a variety of other measures from 1980.[11] The site-specific characteristics include the Hazard Ranking System (HRS) score of the closest site, a dummy variable if the HRS score is missing for that site, and the years elapsed since the closest site was first "discovered" by the EPA for inclusion on the NPL. As every observation (i.e., block group) has a closest NPL site, the HRS variable should say little about the environmental quality of the observation; merely it should indicate the likelihood of that site to be cleaned up (Hird 1990, Layton and Daley 2004). Likewise, the time elapsed since discovery should affect the likelihood of clean-up rather than environmental quality of that neighborhood. This instrumented version, $\hat{\dot{E}}$, replaces $\dot{E}$ in Models 2 and 4.

The concern that $\dot{E}$ may be correlated with other unobservables that belong in the empirical model motivates several additional models. First, Model 1' represents Model 1 except with only the subsample of neighborhoods with NPL sites in 1990. Second, Model 2 duplicates Model 1 except that $\dot{E}$ is replaced by $\hat{\dot{E}}$. Third, the possibility of time-varying hedonic prices (β) is explored in Models 2G and 2EG.[12] Even an instrumented $\hat{\dot{E}}$ might give biased price effects if clean-ups tend to occur in areas that tended to experience changes in price (due to either demand or supply shocks) during the 1990s. Suppose that clean-ups occurred in areas with an NPL site in 1990 or with a rainy climate. Obviously, both attributes are time-invariant and thus cannot explain appreciation rates during the 1990s, unless hedonic prices changed during that time. The assumption of time-invariant effects for levels of E_{t-1} and G are relaxed in order to control for the possibility that clean-ups tended to occur in areas that experienced price changes. Models 2G and 4G include in the price equation the G vector of the natural amenity scale, distance to city center, and MSA-specific distances. Letting prices for these attributes vary controls for the possibility that, for instance, if downtown areas tended to get clean-ups and also tended to see large price increases during the 1990s, then β_E might be biased upwards without controlling for distance-to-downtown. Finally, Models 2EG and 4EG include both the G vector and the E_{t-1}, allowing for the possibility that block-groups with or near NPL sites in 1990 were somehow different than other areas and followed a different price path over the 1990s. Table 2 summarizes the differences between models.

Table 2: Summary of Models

Model	Equation	Sample	Vectors included
1	(2) OLS	Full	$\dot{S}$, $\dot{N}$, $\dot{M}$, $\dot{E}$
1'	(2) OLS	NPL in 1990 only	$\dot{S}$, $\dot{N}$, $\dot{M}$, $\dot{E}$
2	(2) OLS	Full	$\dot{S}$, $\dot{N}$, $\dot{M}$, $\hat{\dot{E}}$
2G	(2) OLS	Full	$\dot{S}$, $\dot{N}$, $\dot{M}$, $\hat{\dot{E}}$, G
2EG	(2) OLS	Full	$\dot{S}$, $\dot{N}$, $\dot{M}$, $\hat{\dot{E}}$, E, G
3	(7) 3SLS	Full	$\dot{S}$, $\dot{N}$, $\dot{M}$, $\dot{E}$
3'	(7) 3SLS	NPL in 1990 only	$\dot{S}$, $\dot{N}$, $\dot{M}$, $\dot{E}$
4	(7) 3SLS	Full	$\dot{S}$, $\dot{N}$, $\dot{M}$, $\hat{\dot{E}}$
4G	(7) 3SLS	Full	$\dot{S}$, $\dot{N}$, $\dot{M}$, $\hat{\dot{E}}$, G
4EG	(7) 3SLS	Full	$\dot{S}$, $\dot{N}$, $\dot{M}$, $\hat{\dot{E}}$, E, G
3a	(7a) 3SLS	Full	$\dot{S}$, $\dot{N}$, $\dot{M}$, $\dot{E}$
3a'	(7a) 3SLS	NPL in 1990 only	$\dot{S}$, $\dot{N}$, $\dot{M}$, $\dot{E}$
4a	(7a) 3SLS	Full	$\dot{S}$, $\dot{N}$, $\dot{M}$, $\hat{\dot{E}}$
4Ga	(7a) 3SLS	Full	$\dot{S}$, $\dot{N}$, $\dot{M}$, $\hat{\dot{E}}$, G
4EGa	(7a) 3SLS	Full	$\dot{S}$, $\dot{N}$, $\dot{M}$, $\hat{\dot{E}}$, E, G

IV. RESULTS

Table 3 summarizes the results of estimating alternative $\dot{P}$ equations based on Equation (2). Table 4 summarizes the results of estimating several alternative specifications of the neighborhood transition and NPL clean-up system in Equation (7). Table 4a does likewise for the alternative system in Equation (7a). (Note that the price equation is the same in both Equation (7) and (7a), leaving OLS estimates of the $\dot{P}$ equation identical.) Model 1 is just a first-differenced price equation, Equation (2), where $\dot{S}$ and $\dot{N}$ are treated as exogenous. Model 3 refers to the system in Equation (7), including the $\dot{P}$ equation and 19 equations for vectors $\dot{S}$ and $\dot{N}$. Models 2 and 4 replicate Models 1 and 3, respectively, except that the instrumented version $\hat{\dot{E}}$ replaces

$\dot{E}$. Models 1’ and 3’ correspond to Models 1 and 3 except they are estimated using the restricted sample of only block groups that contained NPL sites in 1990 or are adjacent to those block groups. This sample restriction reduces the sample size from about 200,000 to about 11,000.

Estimates of Models 1 and 2 in Table 3 show median housing value changing with changes in structural and demographic characteristics of the neighborhood. The initial first-difference price regression offers results generally consistent with expectations and previous literature. The effects of these controls are relatively stable across models in Table 3, especially for the full sample. Newness, rooms, plumbing, income, percent white, education, absence of poverty, and shorter commutes are all associated with higher prices in the OLS model. Model 1’ shows that a somewhat different set of prices may hold in the subsample. Otherwise, the OLS models explain about 22% of the variation in appreciation rates in block groups in fairly standard ways.

[Insert Table 3 about here]

NPL site clean-up is associated with a 3.7% rise in prices in Model 1 (with a 95% confidence interval of 2.1-5.3%). This direct effect is robust to a variety of other controls, removing the MSA-level fixed effects, or even restricting the sample to only block-groups with or near an NPL site 1990. Model 1’ shows a 4.2% direct price effect of clean-ups among the smaller sample. Instrumenting for $\dot{E}$ in Model 2, on the other hand, essentially eliminates the direct price effect. The largest effect is seen in Model 2EG, where the direct price effect is estimated to be 5.8% but with a 95% confidence interval ranging from -3.9% to 15.5%. OLS estimates an effect size of $\dot{E}$ to be about 4%, whether in the full sample or only among areas with or near NPL sites, while the price effect of $\hat{\dot{E}}$ is statistically indistinguishable from zero. Overall, these results are within the range of the findings referenced in Section II.

The direct price effects of clean-ups are estimated in Tables 4 and 4a using the 3SLS approach. Depending on the modeling assumptions and sample, the estimated price effects range from -9.5% to 6.5% in Table 4, but only the estimate in Model 3 is significantly different from zero.[13] The alternative models in Table 4a exhibit a similar range in estimates, although most are closer to zero.[14] The model that includes E_{t-1} proves to be the exception, with the price effect of $\hat{\dot{E}}$ tending to be more negative in Model 4EGa than its counterpart in Model 4EG. In Model 3a, median property values in block groups with or near clean-ups appreciate a statistically insignificant 2.7% faster than other block groups. The instrumented clean-up variable produces very similar direct price effects in Model 4a. Among neighborhoods with or near NPL sites, however, block groups with remediations actually appreciate 1.6% more slowly on average. Only when E_{t-1} and G are included does the direct price effect of instrumented clean-ups become statistically different from zero. In Model 4EGa, block groups experiencing remediations exhibit 20.7% lower appreciation rates than comparable properties.

[Insert Table 4 about here]

[Insert Table 4a about here]

Comparing across Tables 4 and 4a, the results broadly suggest a small and noisy direct price effect from de-listings under the null of exogenous clean-ups, and this effect erodes further when clean-ups are instrumented. Omitting the initial proximity to old hazardous waste sites from the model may upwardly bias the price effect of $\dot{E}$ or $\hat{\dot{E}}$. Apparently those neighborhoods around NPL sites circa 1990 experienced above average appreciation during the 1990s. This weak result, that β_E may be incorrectly attributing some of the price growth in "dirty" neighborhoods to the clean-ups that occurred in *some* of them, is evident from comparing β_E in Models 3 and 3', and 4G and 4EG. The similar pattern of upward bias holds in Table 4a. The bias is so strong in the alternative Model 4EGa that the negative direct price effect actually becomes significant. Otherwise, the evidence in Tables 4 and 4a point to consistently small and insignificant effects of clean-ups.

This result resembles the OLS results presented in Table 3, but it contrasts with much of the hedonic literature on Superfund sites. These single-site studies often find significant price effects on properties even farther away than one or two block groups. Kiel and Williams (2006) are an exception in this literature, where their multi-site approach finds positive price effects of proximity over time for some of their sample. In addition, our results suggest that unobservables associated with the neighborhoods around NPL sites and, in particular, sites experiencing remediations may bias the observed price effects.[15]

The effects of clean-ups on neighborhood composition and housing stock are presented in the bottom panels of Table 4 and Table 4a. In many cases, NPL site clean-ups are associated with significant changes in structural and demographic changes in the neighborhood. In Model 3, $\dot{E}$ predicts newer housing, a larger share of housing with gas or electric heat, and larger shares of small structures and blue-collar workers. In the smaller sample in Model 3', $\dot{E}$ is insignificant in all of the other equations in the system.

Instrumenting for $\dot{E}$ in an auxiliary regression leads to generally larger (and more significant) effects of $\hat{\dot{E}}$ in the other equations in the system. These effects differ depending on whether $E_{t\text{-}1}$ or G are included. Some consistent effects are evident, however. The positive effect of $\hat{\dot{E}}$ on *Years built* and negative effect on *Income* and *Blue collar* appear to be the result of omitting $E_{t\text{-}1}$, a measure of baseline neighborhood environmental quality. Areas with or near NPL sites in 1990 tended to become newer, poorer, and less blue collar over the 1990s, rather than just the areas experiencing remediations. Model 4EG, which includes $E_{t\text{-}1}$ in each of the $\dot{S}$ and $\dot{N}$ equations and uses both $E_{t\text{-}1}$ and G as exogenous instruments in the first stage, finds several significant non-price effects of clean-ups. Neighborhoods with or near a clean-up tend to gain more rooms per housing unit but fewer bedrooms per unit, a larger share of units using utilities for their heat, a larger share of nonwhite residents, a larger share of units occupied by renters, and shorter commutes. As shown in the top panel of Table 4, some of these affected variables are significant in the hedonic price equations; others are not. Interestingly, block-groups nearby NPL deletions during the 1990s did not appear to become much wealthier, more white and educated, and more family friendly as some might expect.

The alternative models in Table 4a show roughly similar effects of clean-ups on neighborhood composition when the effects are significant. With population densities rising and median building age becoming younger, one might conclude that new housing is being built as new residents move into these areas. This provides some evidence for both supply and demand effects of changes in environmental quality.

Although housing markets and residential sorting mechanisms appear responsive to changes in environmental quality, direct price estimates, from neighborhood-level hedonic analyses (as in Models 1 or 2) or from systems models (as in Models 3 or 4), capture only part of the effect of clean-ups on prices. The full price effect of an NPL clean-up can be calculated via equation 9 or 9a. These effects appear in Table 5 along with direct price effects reprinted from Table 4.

In most of the models estimated, the indirect effect goes in the opposite direction of the direct effect. Relative to the standard errors for β_E, these indirect effect sizes are often fairly small. Two exceptions arise in Models 3' and 4EG – both of which control for E_{t-1} either by directly including it or by restricting the sample. Clean-ups in either case do not have significant direct price effects, but they do induce changes in demographics and housing stock such that prices substantially rise. The indirect price effect in Model 3', which uses a sample restriction to mitigate the policy endogeneity, appears substantively large and positive. The net price effect of clean-ups among the eligible neighborhoods is a 3.8% increase, which is a substantial improvement over the 4.3% decrease that would be expected under a clean-up if the neighborhood dynamics were held fixed. Using G as instruments and including G in the price equation in the full sample also obtains positive indirect price effects. On net, the flexible full-sample specification with MSA-specific effects of proximity to city center predicts 7.8% higher property values in block groups with or near NPL deletions. This positive price effect, however, arises through neighborhood change occurring and influencing price. Direct price effects alone tend not to be positive and significant in these models. Overall, the results suggest two things: (1) the direct price effects are not large and positive; and (2) allowing for indirect price effects can yield significant total price effects from clean-up, but much noise remains and the results are highly sensitive to model specification. Similar results hold for the alternative model.[16] The preferred model points to a +7.8% total price effect from clean-ups, with a -9.5% direct and a +17.3% indirect price effect.

The net result is that clean-ups tend not to appear as an amenity, unless the clean-up's induced changes in neighborhood composition are also allowed to affect price. Throughout all of the models estimated, the direct price effect is never significant and positive, and their total price effects are also centered near zero. (If the confidence interval around the full effect estimate was the same size as that of the direct effect, the full effect's confidence interval would contain zero for all models.) With a full set of controls in Models 4EG and 4EGa, even if the direct price effect appears to be negative on average, the indirect effects bring the net price effect to positive through neighborhood change. Much noise surrounds these estimates.

The Models 3' and 3a' offer useful reference points, as they use a sample restriction to severely limit the comparison group. Among these observations eligible for receiving an NPL remediation, the direct price effect appears small, negative, and insignificant, whereas the full price effect appears positive and small (~4%). Though mixed, there is some evidence for positive indirect (not direct) price effects of NPL remediation during the 1990s.

[Insert Table 5 about here]

V. DISCUSSION

In this paper, we consider the price effects of changes in environmental quality in two important dimensions often overlooked in the literature. First, we explicitly model neighborhoods (block groups) as panel data in a first-difference model. This allows for better controls of omitted variables and allows explicit estimation of within-observation covariation in prices and environmental change. Second, we treat important attributes of the neighborhood (*S* and *N*) as simultaneously determined. We estimate the direct and indirect pathways through which changes in environmental quality can affect prices. The evidence suggests that there are weak indirect effects on prices through induced changes in *S* and *N* in this context.

While hedonic prices may be relatively easy to compute, using these estimates as predictions of policy effects requires great care. Hedonic prices derived from variation in environmental quality (*E*) across units are often interpreted as marginal willingness to pay to improve *E*. This marginal price, β, clears the market when households choose among properties with varying environmental quality. Yet, many unobserved attributes of housing likely correlate with *E*. Repeat-sales using panel data can help researchers avoid attributing price effects of these unobservables to policy interventions. More importantly, as the results here suggest, even unbiased estimates of β may be inappropriate for predicting the price effects of a change in *E*. An estimated β that explains between-observation variation in price may be a poor predictor of within-observation price changes in response to changes in *E*. Shocks to *E* may induce shifts in housing and other markets, and the joint determination of several important variables like price and neighborhood composition. An estimator that reflects the partial price effect, holding key neighborhood composition variables fixed, may overlook significant changes in those variables induced by the policy intervention.

In principle, estimating richer models of the joint determination of prices, neighborhood composition, and environmental quality can offer important insight into these indirect effects. This paper estimated those rich models for a major brownfields clean-up program, one which has cost taxpayers roughly $30 billion (Greenstone and Gallagher 2005), and finds mixed evidence of significant direct or indirect price effects. If Superfund has any positive impact on property values, it seems that it must come through induced changes in housing stock and neighborhood composition. How these indirect effects should be used in, say, a cost-benefit analysis depends on the context. If a clean-up attracts housing investment or high-income families, some of that investment and in-migration is coming at the expense of other areas. Thus, these indirect effects should be used judiciously by policy-makers interested in efficiency. More local interests may care less about effects in other areas or markets.

Although evidence on price effects is weak here, estimating the structural models (equation 7) reveals much more information than just price effects. The effects of EPA clean-ups on neighborhood composition, as local housing markets adjust to changes in the urban environment, can be seen in the bottom panels of Tables 4 and 4a. In our preferred model (4EG), remediations tended to attract a rising share of minorities and renters to the neighborhood.[17] Remediations do not explain changes in income, education, or percent children. These empirical results contribute to the growing literature on neighborhood transition and environmental change (e.g., Banzhaf and Walsh 2006, Cameron and McConnaha 2006).

These findings also have important implications for the environmental justice debate. Ringquist (2006) reviews much of the evidence on the spatial correlation of disamenities and demographic groups, both in static and dynamic settings. He finds little evidence in the literature that observed inequitable distributions result from sorting induced by environmental change, though his review does not consider the recent research mentioned here (e.g., Banzhaf and Walsh 2006). The results in the bottom panel of Table 4 suggest that some sorting does indeed take place following environmental change. Remediations precede demographic shifts in neighborhoods, although the shifts are small and perhaps act in unexpected directions. After a site is removed from the NPL, the share of minority residents nearby increases. This suggests that remediating brownfields may pave the way for certain demographic groups to move into newly cleaned areas. Conversely, these effects also imply that attempts to catalyze urban renewal with brownfield remediation may not lead to the intended outcomes. As is common, individual and market behavior can undermine the best of policy intentions. Original residents may depart and new residents may arrive. Moreover, their ability to capture any property value appreciation depends on ownership, and the results in Table 4a hint at new rental housing following remediations.

The present research invites further inquiry into simultaneous neighborhood and environmental change. A more robust system would better control for endogeneity in listing and remediation of NPL sites. A general equilibrium approach might also model other important markets, such as the labor market, to fully assess the expected price changes associated with remediation. Recent applications to air quality (e.g., Bayer et al. 2003, Sieg et al. 2004) demonstrate the utility of general equilibrium models in examining joint environmental and neighborhood change. Certainly micro-level data would allow for more useful estimates and validation of our findings in local markets.

Whether price effects of NPL sites vary across sites or metropolitan areas, perhaps using a random coefficients framework, warrants additional attention following on Kiel and Williams (2006).

More generally, the approach taken here can be extended to other contexts to enrich the use of hedonic estimates to guide and evaluate public policy. The evidence on Superfund clean-ups on property values is weak and inconsistent – adding to recent findings by Greenstone and Gallagher (2005) and Kiel and Williams (2006). More than reinforcing these studies, our approach models a large and complex set of relationships extending well beyond price effects. This promises a more detailed picture of the neighborhood dynamics following environmental change. Although Superfund clean-ups had relatively small and inconsistent impacts on many variables, as with price, this sort of approach may yield great insights in other applications. For example, Shapiro's (2005) recent analysis of air toxics models risk changes as a function of static demographic variables. As this paper demonstrates, a simultaneous approach with changes in demographic variables is likely to produce different, and much richer, results. Finally, the sensitivity of our results to different assumptions about policy endogeneity indicates that this is no small concern in the case of Superfund. How listings or clean-ups are assigned is crucial to both identifying the many impacts of the policy and generalizing from observed impacts at one site to another.

References

Arellano, Manuel (1989) "A Note on the Anderson-Hsiao Estimator for Panel Data," *Economics Letters*, vol. 31, no. 4, pp. 337-341.

Baden, Brett and Don Coursey (2002) "The Locality of Waste Sites Within the City of Chicago: A Demographic, Social, and Economic Analysis," *Resource and Energy Economics*, vol. 24, pp. 53-93.

Banzhaf, Spencer and Randall Walsh (2006) "Do People Vote with Their Feet? An Empirical Test of Environmental Gentrification." *Resources for the Future Discussion Paper 06-10*.

Bayer, Patrick, Nathaniel Keohane, and Christopher Timmins (2006) "Migration and Hedonic Valuation: The Case of Air Quality," NBER Working Paper No. 12106.

Been, Vicki (1997) "Coming to the Nuisance or Going to the Barrios? A Longitudinal Analysis of Environmental Justice Claims," vol . 24, no. 1, *Ecology Law Quarterly*, pp. 1-56.

Cameron, Trudy Ann and Ian T. McConnaha (2006) "Evidence of Environmental Migration," *Land Economics*, vol. 82, no. 2, pp. 273-290.

Clark, David E. and Leslie A. Nieves (1994) "An Interregional Hedonic Analysis of Noxious Facility Impacts on Local Wages and Property Values," *Journal of Environmental Economics and Management*, vol. 27, pp. 235-253.

Coulton, C., T. Cook, et al. (2004). "Aggregation Issues in Neighborhood Research: A Comparison of Several Levels of Census Geography and Resident Defined Neighborhoods." APPAM Fall Research Conference, Atlanta, GA.

Dale, Larry, James C. Murdoch, Mark A. Thayer, and Paul A. Waddell, (1999) "Do Property Values Rebound from Environmental Stigmas? Evidence from Dallas" *Land Economics*, vol. 75, no. 2, pp. 311-326.

Deaton, B. James and John P. Hoehn (2004) "Hedonic Analysis of Hazardous Waste Sites in the Presence of Other Urban Disamenities," *Environmental Science and Policy*, vol. 7: 499-508.

Diamond, Douglas B. Jr. and George S. Tolley (1982) "The Economic Roles of Urban Amenities," in Diamond, Douglas B Jr., and George S. Tolley, eds., *The Economics of Urban Amenities*, (New York, New York: Academic Press), pp. 3-54.

Ekeland, Ivar, James J. Heckman, and Lars Neshiem, (2002) "Identifying Hedonic Models," *American Economic Review*, vol. 92, no. 2, pp. 304-309.

Freeman, A. Myrick III, (1979) "The Hedonic Price Approach to Measuring Demand for Neighborhood Characteristics," in Segal, David, ed. *The Economics of Neighborhood* (New York, Academic Press), pp. 191-217.

Gayer, Ted (2000) "Neighborhood Demographics ad the Distribution of Hazardous Waste Risks: An Instrumental Variables Estimation," *Journal of Regulatory Economics*, vol 17, no. 2, pp. 131-155.

Gayer, Ted, James T. Hamilton, and W. Kip Viscusi (2000) "Private Values of Risk Tradeoffs at Superfund Sites: Housing Market Evidence on Learning about Risk," *Review of Economics and Statistics*, vol. 82, no. 3, pp.439-451.

Goodman, Allan C. (1977) "A Comparison of Block Group and Census Tract Data in a Hedonic Housing Price Model," *Land Economics*, vol. 53, no 4, pp. 483-487.

Greenberg, M. and J. Hughes (1992) “The Impact of Hazardous Waste Superfund Sites on the Value of Houses Sold in New Jersey,” *Annals of Regional Science*, vol. 26, pp. 147-153.

Greenstone, Michael and Justin Gallagher (2005) “Does Hazardous Waste Matter? Evidence from the Housing Market and the Superfund Program,” NBER Working Paper No. 11790.

Gupta, Shreekant, George van Houten, and Maureen Cropper (1996) “Paying for Permanence: An Economic Analysis of EPA's Cleanup Decisions at Superfund Sites,” *RAND Journal of Economics*, vol. 27, no. 3, pp. 563-582.

Hamilton, James T. (1995) “Testing for Environmental Racism: Prejudice, Profits, Political Power,” *Journal of Policy Analysis and Management*, vol. 14, no. 1, pp. 107-132.

Hird, John A (1990) “Superfund Expenditures and Cleanup Priorities: Distributive Politics of the Public Interest?” *Journal of Policy Analysis and Management*, vol. 9, no. 4, pp. 455-483.

Hird, John A. (1993) “Environmental Policy and Equity: The Case of Superfund,” *Journal of Policy Analysis and Management*, vol. 12, no. 2, pp. 323-343.

Hird, John A. (1994) *Superfund: The Political Economy of Environmental Risk*, Baltimore: Johns Hopkins University Press

Ketkar, Kusum (1992) “Hazardous Waste Sites and Property Values in the State of New Jersey,” *Applied Economics,* vol. 25, pp. 647-659.

Kiel, Katherine A. (1995) "Measuring the Impact of the Discovery and Cleaning of Identified Hazardous Waste Sites on House Values," *Land Economics*, vol. 71, no. 4, pp. 428-435.

Kiel, Katherine A. and Michael Williams (2005) "The Impact of Superfund Sites on Local Property Values: Are All Sites the Same?" College of the Holy Cross, Department of Economics Faculty Research Series, Working Paper 05-05.

Kiel, Katherine A. and Jeffrey E. Zabel (2001) "Estimating the Economic Benefits of Cleaning Up Superfund Sites: The Case of Massachusetts," *Journal of Real Estate and Finance Economics*, vol 22, no. 2-3, pp. 163-184.

Kohlhase, Janet E. (1991) "The Impact of Toxic Waste Sites on Housing Values" *Journal of Urban Economics*, vol. 30, no. 1, pp. 1-26.

McCluskey, Jill J. and Gordon C. Rausser (2003) "Stigmatized Asset Value: Is it Temporary of Long-Term?" *Review of Economics and Statistics*, vol. 85, no. 2, pp. 276-285.

Mendelsohn, Robert, Daniel Hellerstein, Michael Huguenin, Robert Unsworth, and Richard Brazee (1992) "Measuring Hazardous Waste Damages with Panel Models," *Journal of Environmental Economics and Management*, vol. 22, pp. 259-271.

Michaels, R. Gregory and V. Kerry Smith (1990) "Market Segmentation and Valuing Amenities With Hedonic Models: The Case of Hazardous Waste Sites," *Journal of Urban Economics*, vol. 28, pp. 223-242.

National Atlas of the United States (2004) *Cities and Towns of the United States*, National Atlas of the United States, Reston, VA. Last accessed online on 10 August 2005 at: http://nationalatlas.gov/atlasftp.html

Nelson, Jon P. (1979) "Airport Noise, Location Rent, and the Market for Residential Amenities," *Journal of Environmental Economics and Management*, vol. 6, pp. 320-331.

O'Byrne, Patricia Habuda, Jon P. Nelson, and Joseph J. Seneca, (1985) "Housing Values, Census Estimates, and the Environmental Cost of Airport Noise: A Case Study of Atlanta," *Journal of Environmental Economics and Management*, vol. 12, no. 2, pp. 169-178.

Ringquist, Evan (2006) "Environmental Justice: Normative Concerns, Empirical Evidence, and Government Action." In Norman J. Vig and Michael E. Kraft, eds. *Environmental Policy: 6th Edition*, Washington, DC: CQ Press.

Rosen, Sherwin (1974) "Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition," *Journal of Political Economy*, vol. 82, pp.34-55.

Shapiro, Marc D. (2005) "Equity and Information: Information Regulation, Environmental Justice, and Risks from Toxic Chemicals," *Journal of Policy Analysis and Management*, vol. 24, no. 2, pp. 373-398.

Shultz, Steven D. and David A. King (2001) "The Use of Census Data for Hedonic Price Estimates of Open-Space Amenities and Land Use," *Journal of Real Estate Finance and Economics* vol. 22, no. 2/3, 239-252.

Sieg, Holger, V. Kerry Smith, H. Spencer Banzhaf, and Randall P. Walsh (2004) "Estimating the General Equilibrium Benefits of Large Changes in Spatially

Delineated Public Goods," *International Economic Review*, vol. 45, no. 4, pp. 1047-1077.

Staiger, Douglas and James H. Stock (1997) "Instrumental Variables Regression with Weak Instruments," *Econometrica*, vol. 65, no. 3, pp. 557-586.

US EPA (2004) "Superfund National Accomplishments Summary Fiscal Year 2004," http://www.epa.gov/superfund/action/process/numbers04.htm

US EPA (2000) "Superfund 20th Anniversary Report"

USDA ERS (1999) *Natural Amenities Scale* data file. Last accessed online 10 August 2005 at: http://www.ers.usda.gov/Data/NaturalAmenities/

Zabel, Jeffrey E. (1999) "Controlling for Quality in Price Indices," *Journal of Real Estate Finance and Economics*, vol. 19, no. 3, 223-241.

Endnotes

[1] Chattopadhyay et al. (2005), Clark and Nieves (1994), Dale et al. (1999), Deaton and Hoehn (2004), Gayer et al. (2000), Ketkar (1992), Kiel (1995), Kiel and Williams (2007), Kiel and Zabel (2001), Kohlhase (1991), McCluskey and Rausser (2003), Mendelsohn et al. (1992) and Michaels and Smith (1990).

[2] This approach is essentially a partial adjustment model for S. Let S^* be the (unobserved) equilibrium and $S^*=f(P, N, E, M, G)+\theta$, such that the observed $S_t - S_{t-1} = \delta[S_t^* - S_{t-1}]$. Thus, Equation (3) in levels can be written as $S_t = (1-\delta)S_{t-1} + \delta f(.) + \delta\theta_t$. Equation (4) is merely the first-differencing of this equation for an f linear in its arguments.

[3] The partial correlation of $\dot{E}$ and $\dot{S}$ would depend on whether housing and environmental quality are complements or substitutes.

[4] The assumption of no endogeneity in (4) would be represented by zeros in the off-diagonal elements of the coefficient matrix except in the first row.

[5] The total effect of changes in E on appreciation rates is given as Equation (8a):

$$\text{(8a)} \qquad \frac{d\dot{P}}{d\dot{E}} = \frac{\beta_E + \beta_S\gamma_E + \beta_N\delta_E + \beta_S\gamma_N\delta_E + \beta_N\delta_S\gamma_E - \beta_E\gamma_N\delta_S}{1-\beta_S\left(\gamma_N\delta_P + \gamma_P\right) - \beta_N\left(\delta_S\gamma_P + \delta_P\right) - \gamma_N\delta_S}$$

[6] A national sample, while arguably lumping multiple markets together, moves away from the single-market analyses that dominate this literature and provide little more than case studies. This approach has both strengths and weaknesses. In this, we follow the leads of Greenstone and Gallagher (2005), Bayer et al. (2005), and Kiel and Williams (2006) in using a large geographic sample. The nature of the problem, sorting within and across housing markets, necessitates exploring multiple regions simultaneously. Ekeland et al.'s (2002) skepticism about justifying a segmented markets approach provides additional motivation for including a national sample in the same estimation. In all of our models, we allow for separate market trends by estimating the price equation with metropolitan-area fixed effects. Further, in some models, we estimate metropolitan-area-specific distance gradients as well. Regardless, the average effect across all clean-ups is of primary interest here, and this warrants the use of a national sample.

[7] See Shultz and King (2001) for additional review of the use of aggregated census data in hedonics. Greenstone and Gallagher (2005) use a similar data set for similar purposes, although they use the larger geography of the census tract.

[8] Up to this point, we have been assuming that environmental quality is exogenous. This simplifies the discussion. We will discuss this assumption, and relax it, later in the paper.

[9] All models estimated here control for MSA-level fixed effects in all equations in the system. It should be emphasized that, when it is specified that *G* is included in the model by relaxing the assumption of fixed prices, this relaxation is applied only to the price equation. The *G* vector enters the price equation only, not the other equations for $\dot{S}$ and $\dot{N}$ in Equation (7) or (7a).

[10] There are a few exceptions because of changes in the census. First is the percent of units in structures containing four or fewer units. Because of coding changes between census years, the twice-lagged value and once-lagged difference are not available. Instead we use the twice-lagged level and once-lagged difference of the percent of housing units in structures with *nine* or less housing units. Second, because of the changes in definition of race and ethnicity between the 1990 and 2000 censuses, the percent white variable might not be precisely comparable across these decades. Also, while in 1990 and 2000 the data on education are reported for person's aged 25 or older, in 1980 they are reported for those 18 and over, so again, the lagged variable is not exactly identical in definition to the endogenous variables.

[11] Other 1980 measures include: median household income; per-capita income; median year built of housing units; median rent; median housing value; percents of housing units vacant, with complete plumbing, with complete kitchens, with gas or electric heat, not sharing structure with other units, sharing structure with 9 or less units, boarded up, with a telephone, with at least 2 bathrooms, with no air conditioner, and with central air conditioner; percents of population that is white, aged 18 or less, aged 65 or more, and with income < 1.5 times poverty line; percent of households that are renters; percent of household with no vehicles; percent of adult population with at least bachelor's degree; percent of population aged at least 25 with no HS degree; average commute travel time (in minutes); number of bedrooms per housing unit; number of rooms per housing unit; housing units per mile^2; population per mile^2; population per housing unit; population per room; percent of workers employed in "industrial" sectors; percent of vacant residences that are boarded up; percent of occupied housing units with

telephones; ln(median housing value); ln(median contract rent); ln(median household income); ln(per-capita income); population density^2; and MSA population.

[12] Ideally, the possibility of time-varying β_S and β_N could also be explored, but including additional endogenous regressors (S_{t-1} and N_{t-1}) in the price equation would leave the system underidentified without resorting to additional arbitrary exclusion restrictions. The results thus depend on the assumption that significant implicit price changes do not occur if that attribute is correlated with remediation.

[13] The instruments used to identify the system in Equation (7) work well, as expected. Staiger and Stock (1997) suggest first-stage F-statistics as indicators of instrument strength. The first-stage regressions in the 3SLS estimation for Models 3 and 4 all yield F-statistics greater than 20 for 38 out of 39 equations, with most F-statistics well exceeding 100. Another concern is that the exclusion restrictions overidentify the system. In Models 3 and 4, the price equation has 19 endogenous variables and 20 exclusion restrictions. The Sargan test statistics (distributed χ^2 with 1 degree of freedom) for the $\dot{P}$ equations, estimating in 2SLS, in Models 3 and 4 range between 57 and 67. These large statistics give some concern that the system is overidentified. Consequently, using this instrument set cannot rule out the endogeneity of $\dot{E}$ in the price equation. It does, however, appear to be exogenous in many of the other equations in the system. Among a restricted sample of neighborhoods with or near NPL sites in 1990, the exogeneity of $\dot{E}$ cannot be rejected in the price equation for Model 3' based on the C-statistic in 2SLS. Overall, the diagnostic statistics for the system estimations in Models 3 – 4 point to strong instruments and only moderate concerns about overidentification.

[14] The system diagnostics in the alternative models (Models 3a – 4EGa) are similar to the diagnostics for the standard models (Models 3 – 4EG). The instruments all appear quite strong based on first-stage F-statistics. The overidentification tests, based on the Sargan statistic calculated from 2SLS estimation, are the same for the price equation – where models 3 and 3a are identical – and are zero for the *S* and *N* equations because they are just-identified. As in Model 3', however, restricting the sample size to only block groups with or near NPL sites in 1990 in Model 3a' appears to solve the overidentification problem. With the Sargan statistic for the price equation in 3a' below 0.50, the small C-statistic is also consistent with the exogeneity of $\dot{E}$ in the price equation of model 3a'.

[15] Incidentally, this result likely does not owe to the choice of comparison group. When the sample is restricted to only the 10,645 neighborhoods with or near NPL sites, the direct price effect of Model 4EG is -0.9836 (z-statistic = -3.72). Clean-ups adversely affect property values even more strongly among the subsample. This result suggests a possible explanation for the negative price effect: the hazardous waste site is a disamenity that harms property values, inclusion in Superfund improves property values, and de-listing it removes the government support without removing the disamenity associated with the site. This may be because the market valued the eyesore more than the health risk, and the EPA only remediated the latter despite high hopes about the former. The negative price effect of de-listing thus reflects, perhaps, a sense of abandonment of the site or neighborhood and only the stigma remains. Perhaps with a longer time-lag after clean-up the negative price effects will dissipate and even become positive. Among these block-groups enjoying a de-listing, the median years passed since deletion is only 2.76.

[16] Large and positive indirect effects are observed for Model 4EGa. The magnitude of the indirect effects in Model 4EGa is comparable to its corresponding indirect effects in Model 4EG. The total price effect in Model 4EGa is estimated to be 4.5%, which is close to the 4.0% estimate from Model 3a' and roughly similar to the 7.8% effect from Model 4EG. Again, as before, the alternative model does not tend to yield positive and significant direct price effects. The full effects in Models 3, 3', 4G, and 4EG are roughly consistent with the corresponding estimates in the alternative model. The most glaring difference arises in comparing Models 4 and 4a. Both exhibit small, positive, and insignificant direct price effects, but the indirect effect is -9.8% in Model 4 and is 16.0% in Model 4a. Including $\dot{P}$ in the non-price equations, without controlling for E_{t-1} or G, greatly affects the results. Overall, similar conclusions hold in the alternative model: positive direct price effects are not observed; and significant indirect price effects can be observed, but they are very sensitive to model specification.

[17] These effects are a bit stronger in Model 4EGa. Yet under the assumption that clean-ups are exogenous in Model 3 or Model 3', clean-ups appear largely orthogonal to demographic trends.

Table 3: OLS models of $\dot{P}$ equation

Model	1	1'	2	2G	2EG
Vectors included	$\dot{S}, \dot{N}, \dot{M}, \dot{E}$	$\dot{S}, \dot{N}, \dot{M}, \dot{E}$	$\dot{S}, \dot{N}, \dot{M}, \hat{\dot{E}}$	$\dot{S}, \dot{N}, \dot{M}, \hat{\dot{E}}, G$	$\dot{S}, \dot{N}, \dot{M}, \hat{\dot{E}}, E, G$
N:	198625	10863	196096	194992	194992
First-differenced variables:	β	β	β	β	β
Year built	0.0013***	0.0014***	0.0013***	0.0013***	0.0013***
Rooms	0.1275***	0.1250***	0.1272***	0.1066***	0.1066***
Utility heat (%)	0.1927***	0.0956**	0.1997***	0.2112***	0.2110***
Housing density[a]	-0.0026	-0.0135	-0.0025	-0.0004	-0.0004
Small structures (%)	-0.1525***	-0.0171	-0.1494***	-0.1522***	-0.1522***
Plumbing (%)	0.2738***	0.1389	0.2640***	0.2039***	0.2039***
Bedrooms	-0.0729***	-0.0647**	-0.0746***	-0.0472***	-0.0472***
Solo unit (%)	-0.0998***	-0.0420	-0.1037***	-0.0743***	-0.0744***
Income	0.1739***	0.1320***	0.1755***	0.1615***	0.1615***
White (%)	0.1855***	0.1058***	0.1864***	0.1605***	0.1605***
College (%)	0.0915***	0.1614**	0.0938***	0.1349***	0.1351***
Poor (%)	-0.1210***	-0.1005*	-0.1255***	-0.0973***	-0.0972***
Blue collar (%)	0.0833***	0.2120***	0.0809***	0.0400***	0.0400***
Renter (%)	0.0608***	0.0782	0.0609***	0.0527***	0.0525***
Children (%)	-0.0224	-0.1625**	-0.0166	-0.0016	-0.0016
Commute	-0.0003*	-0.0015**	-0.0003**	-0.0004***	-0.0004**
No vehicle (%)	-0.0646***	0.1696**	-0.0639***	-0.0784***	-0.0782***
Population density[a]	0.0001	0.0094*	0.0001	-0.0010	-0.0010
Household size	0.0002	-0.0039**	0.0002	0.0002	0.0002
$\dot{M}$	Yes	Yes	Yes	Yes	Yes
NPL in 1990, own/adjacent					0.0087
Clean-up, own/adjacent (β_E)	0.0368***	0.0424***	0.0065	-0.0057	0.0580
Natural amenity scale				-0.0113***	-0.0113***
MSA & distance interactions				Yes	Yes
constant	-0.0001	0.0015	-0.0005	0.0084	0.0084
R^2	0.1911	0.2322	0.1939	0.2234	0.2252

Table 4: Results for $\dot{P}$ equation and selected results for other equations

Model	**3**	**3’**	**4**	**4G**	**4EG**
Vectors included:	$\dot{S}, \dot{N}, \dot{M}, \dot{E}$	$\dot{S}, \dot{N}, \dot{M}, \dot{E}$	$\dot{S}, \dot{N}, \dot{M}, \hat{\dot{E}}$	$\dot{S}, \dot{N}, \dot{M}, \hat{\dot{E}}, G$	$\dot{S}, \dot{N}, \dot{M}, \hat{\dot{E}}, E_{t-1}, G$
N:	197050	10779	195293	194992	194992
First-differenced variables:	β	β	β	β	β
Year built	0.0059***	0.0078***	0.0045***	0.0048***	0.0044***
Rooms	0.0235	1.5874***	0.0046	-0.0695**	-0.0730**
Utility heat (%)	-0.4397***	0.2729	-0.4400***	-1.0336***	-1.0198***
Housing density[a]	-0.0138	2.3110***	-0.0021	-0.0437**	-0.0414**
Small structures (%)	1.3312***	4.9288***	1.8352***	3.3732	3.3777***
Plumbing (%)	10.0976***	-32.0523***	9.0514***	-1.0000	0.1434
Bedrooms	-0.9888***	-1.1482*	-0.8858***	0.1641***	0.1328
Solo unit (%)	-4.5379***	0.7530	-4.7393***	-6.9482**	-6.9409***
Income	0.9667***	-4.2801***	1.1223***	0.2351***	0.2776***
White (%)	-0.8743***	-3.1763***	-0.9003***	-0.6964	-0.6813***
College (%)	-0.3866**	1.4409	-0.2206	-0.0375***	-0.1230
Poor (%)	1.5293***	-13.4137***	1.4343***	-1.6301	-1.3677***
Blue collar (%)	1.1617***	0.6513	0.8935***	0.2590	0.4109**
Renter (%)	1.1874***	5.1639**	1.2317***	0.1673*	0.2262
Children (%)	-3.7740***	-4.7040	-3.6140***	-0.8175***	-0.8780**
Commute	0.0191***	0.0342	0.0212***	0.0518***	0.0513***
No vehicle (%)	-3.1385***	-1.7474	-2.7322***	-2.4506***	-2.6370***
Population density[a]	0.0383***	-0.5593**	0.0331***	0.0942***	0.0951***
Household size	0.0896***	0.0127	0.0777***	-0.0569***	-0.0602***
NPL in 1990, own/adjacent					0.0088
Clean-up, own/adjacent (β_E)	0.0318*	-0.0426	0.0652	-0.0473	-0.0954
constant	0.0114***	0.0076	0.0101***	0.0058	0.0050
Dependent variable:	Partial effect of “Clean-up in or adjacent” by equation, i.e., γ_E or δ_E				
Year built	2.5461**	-1.0998	17.6817***	6.9839***	0.9461
Rooms	0.0393	0.0273	0.1494	0.1149***	0.2146***
Utility heat (%)	0.0288***	-0.0132	0.0962***	0.0449***	0.0933***
Housing density	7.7698	4.0814	-264.906	-155.829***	-64.843
Small structures (%)	0.0067**	0.0119	0.0467***	-0.0016	0.0150
Plumbing (%)	-0.0013	-0.0018	-0.0133***	-0.0121***	-0.0011
Bedrooms	-0.0253	0.0248	-0.1626*	-0.0300	-0.1013***
Solo unit (%)	-0.0075	0.0006	-0.0306	0.0005	0.0004
Income	-0.0155	0.0114	-0.1115**	-0.0695***	-0.0241
White (%)	-0.0245	-0.0072	-0.4895***	-0.0236	-0.0465*
College (%)	-0.0262	0.0069	-2.0040***	0.0158	-0.0103
Poor (%)	-0.0004	-0.0041	0.0110	-0.0262***	-0.0005
Blue collar (%)	-0.0083*	0.0070	-0.0448***	-0.0174*	-0.0144
Renter (%)	0.0085	0.0049	0.0108	0.0149	0.0338*
Children (%)	-0.0048	-0.0011	0.0374	0.0081	-0.0071
Commute	-0.4155	0.1577	-3.1187***	-0.7774	-1.9675**
No vehicle (%)	0.0017	-0.0028	0.0486***	0.0118	-0.0136
Population density	2.7180	-48.4256	134.3119	234.5556	92.8298
Household size	0.0295	-0.0197	-0.6007	-0.1538	-0.5178

[a] measured as 1000s/mi^2

***, **, * for p<0.01, <0.05, <0.10, respectively

Table 4a: Results for $\dot{P}$ equation and selected results for other equations, alternative model

Model	3a	3a’	4a	4Ga	4EGa
Vectors included:	$\dot{S}$, $\dot{N}$, $\dot{M}$, $\dot{E}$	$\dot{S}$, $\dot{N}$, $\dot{M}$, $\dot{E}$	$\dot{S}$, $\dot{N}$, $\dot{M}$, $\hat{\dot{E}}$	$\dot{S}$, $\dot{N}$, $\dot{M}$, $\hat{\dot{E}}, G$	$\dot{S}$, $\dot{N}$, $\dot{M}$, $\hat{\dot{E}}, E_{t-1}, G$
N:	197050	10779	196096	194992	194992
First-differenced variables:	β	β	β	β	β
Year built	0.0043***	0.0058*	0.0034***	0.0052***	0.0049***
Rooms	0.0830	0.9428**	0.0630	0.2086***	0.2062***
Utility heat (%)	-0.3286***	-0.3296	-0.3351***	-0.8062***	-0.7859***
Housing density[a]	0.0130	1.4698	0.0219	-0.0989***	-0.0957***
Small structures (%)	2.0775***	2.4964	1.8635***	3.3629***	3.3891***
Plumbing (%)	3.4392	-18.3533*	4.0472	-0.7552	0.4104
Bedrooms	-0.7231***	-0.7002	-0.7151***	-0.4023***	-0.4428***
Solo unit (%)	-3.9948***	-1.7865	-3.9726***	-5.3543***	-5.3370***
Income	0.9474***	-1.9055	0.9763***	-0.0561	-0.0046
White (%)	-0.9552***	-1.5075	-0.9042***	-0.8585***	-0.8470***
College (%)	-0.2788	-1.1039	-0.3456*	0.0429	-0.0329
Poor (%)	-0.2505	-7.6720*	-0.1822	-2.5823***	-2.3059***
Blue collar (%)	0.4034	0.6094	0.4422	0.7622***	0.9162***
Renter (%)	0.5873**	0.3469	0.5085*	0.7643***	0.8538***
Children (%)	-1.5207***	0.1076	-1.4861***	-1.1289***	-1.2416***
Commute	0.0067	-0.0102	0.0061	0.0172***	0.0160***
No vehicle (%)	-1.5077***	-0.1889	-1.4797***	-2.5275***	-2.7198***
Population density[a]	0.0096	-0.3268	0.0125	0.1068***	0.1074***
Household size	0.0288	0.0093	0.0306	-0.0638***	-0.0644***
NPL in 1990, own/adjacent					0.0281***
Clean-up, own/adjacent (β_E)	0.0270	-0.0159	0.0191	-0.0181	-0.2073***
constant	-0.0055**	0.0089	0.0054**	-0.0044	-0.0048
Dependent variable:	Partial effect of “Clean-up in or adjacent” by equation, i.e., γ_E or δ_E				
Year built	0.1261	2.0354	10.5347***	6.0890***	4.3927*
Rooms	1.3197**	0.0169	3.4809*	0.1063***	0.2507***
Utility heat (%)	0.0353***	0.0400	0.0963***	0.0448***	0.0940***
Housing density	-336.897	10.0214	422.8006	-146.708***	-100.9781
Small structures (%)	-0.0114	0.0069	-0.0037	-0.0024	0.0255***
Plumbing (%)	-0.0073	-0.0010	-0.0045	-0.0110***	0.0036
Bedrooms	0.0090	-0.0233	-0.0374	-0.0232	-0.1300***
Solo unit (%)	0.0030	-0.0068	-0.0056	0.0027	-0.0194*
Income	-0.0236	0.0927	-0.0728	-0.0643***	-0.0076
White (%)	-0.0132	0.0002	-0.2422***	-0.0153	-0.0638**
College (%)	-0.6789	-0.0114	-2.2250	0.0172*	-0.0160
Poor (%)	-0.0462**	-0.0021	-0.1754**	-0.0213***	-0.0029
Blue collar (%)	0.0309	-0.3348	0.1945*	-0.0138	0.0103
Renter (%)	-0.0078	0.0327	-0.0072	0.0135	0.0465**
Children (%)	0.0563	0.0200	0.7671*	0.0083	-0.0158
Commute	0.4030	3.1192	-2.8637	-0.7646	-1.8354**
No vehicle (%)	-0.0544***	0.8093	0.1432***	0.0093	-0.0316**
Population density	2507.145	-48.7680	5866.63	188.3875	677.8631*
Household size	-0.0396	1.3606	-0.9646	-0.1147	-0.3077

[a] measured as 1000s/mi^2
***, **, * for p<0.01, <0.05, <0.10, respectively

Table 5: Summary of direct and full price effects of clean-up actions

Model	Vectors included	$\dot{E}$	Direct effect $\partial P/\partial E$	Indirect effect	Full effect dP/dE
3	$\dot{S}, \dot{N}, \dot{M}, \dot{E}$	clean-up in own or adjacent block group	0.0318	-0.0165	0.0153
3'	$\dot{S}, \dot{N}, \dot{M}, \dot{E}$	clean-up in own or adjacent block group	-0.0426	0.0802	0.0376##
4	$\dot{S}, \dot{N}, \dot{M}, \hat{\dot{E}}$	instrumented clean-up in own or adjacent block group	0.0652	-0.0976	-0.0324
4G	$\dot{S}, \dot{N}, \dot{M}, \hat{\dot{E}}, G$	instrumented clean-up in own or adjacent block group	-0.0473	0.0851	0.0378
4EG	$\dot{S}, \dot{N}, \dot{M}, \hat{\dot{E}}, E_{t-1}, G$	instrumented clean-up in own or adjacent block group	-0.0954	0.1734	0.0780##
		Alternative model			
3a	$\dot{S}, \dot{N}, \dot{M}, \dot{E}$	clean-up in own or adjacent block group	0.0270	0.0284	0.0554
3a'	$\dot{S}, \dot{N}, \dot{M}, \dot{E}$	clean-up in own or adjacent block group	-0.0159	0.0556	0.0397
4a	$\dot{S}, \dot{N}, \dot{M}, \hat{\dot{E}}$	instrumented clean-up in own or adjacent block group	0.0191	0.1600	0.1791##
4Ga	$\dot{S}, \dot{N}, \dot{M}, \hat{\dot{E}}, G$	instrumented clean-up in own or adjacent block group	-0.0181	0.0348	0.0167
4EGa	$\dot{S}, \dot{N}, \dot{M}, \hat{\dot{E}}, E_{t-1}, G$	instrumented clean-up in own or adjacent block group	-0.2073**	0.2523	0.0450##

** indicates significance at the 5% level. ## indicates full effect is outside of the direct effect's 95% confidence interval.